

Phylogenetic shadowing:

**A Novel Tool To Identify Functional
Regions in the Human Genome**

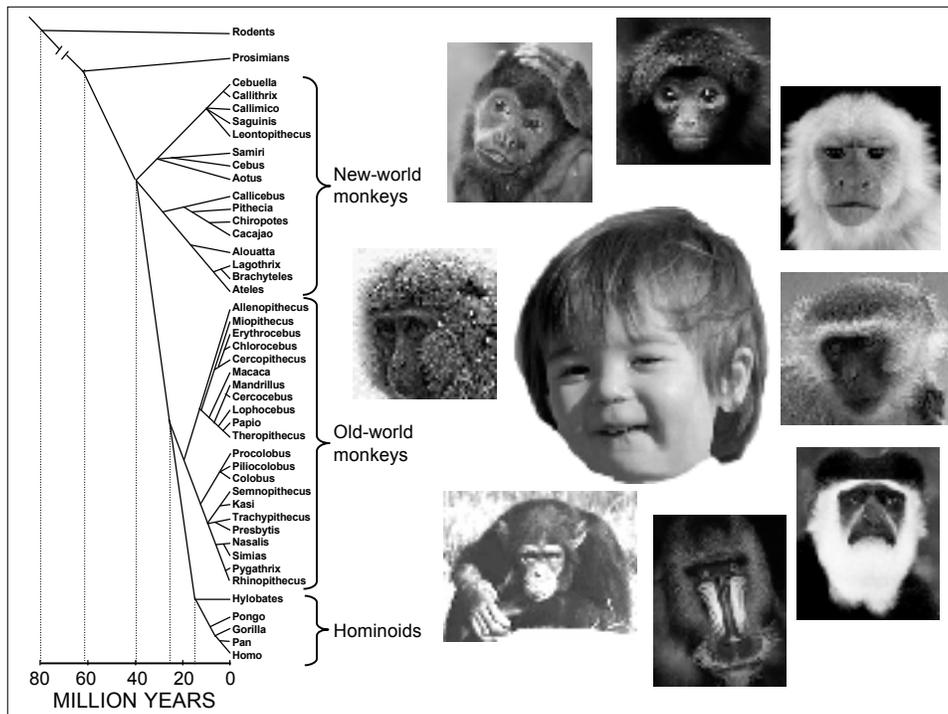
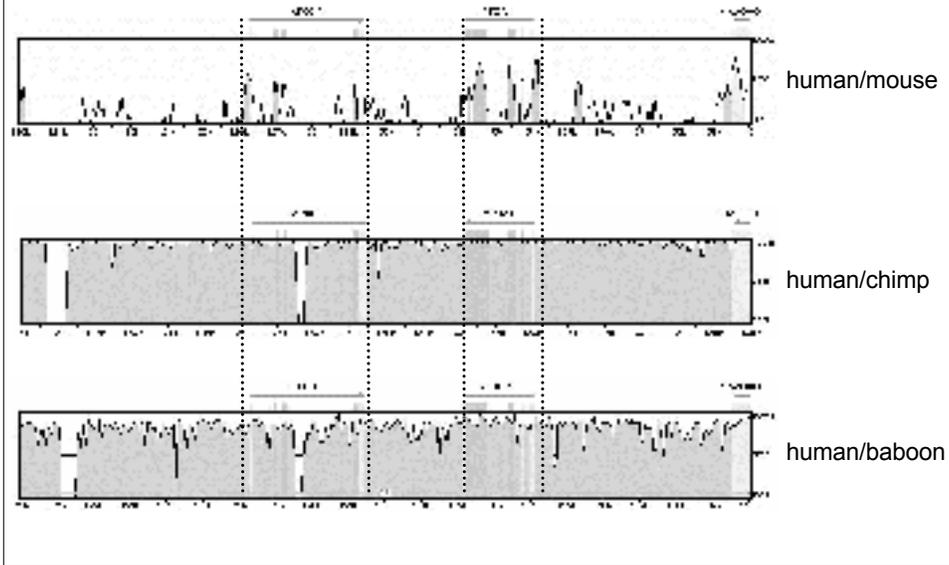
Approach to the Identification of Primate- Specific Regulatory Elements



VS



Pairwise species comparison



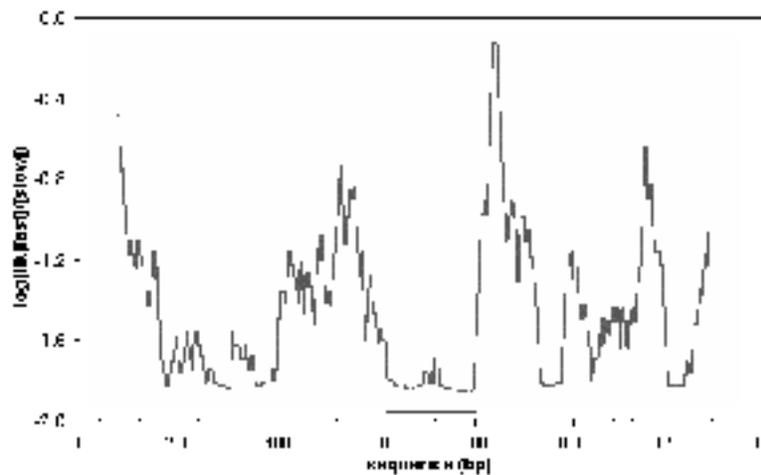
1. PROOF-OF-PRINCIPLE:

**Identification of intron-exon boundaries
of 4 genes with known boundaries**

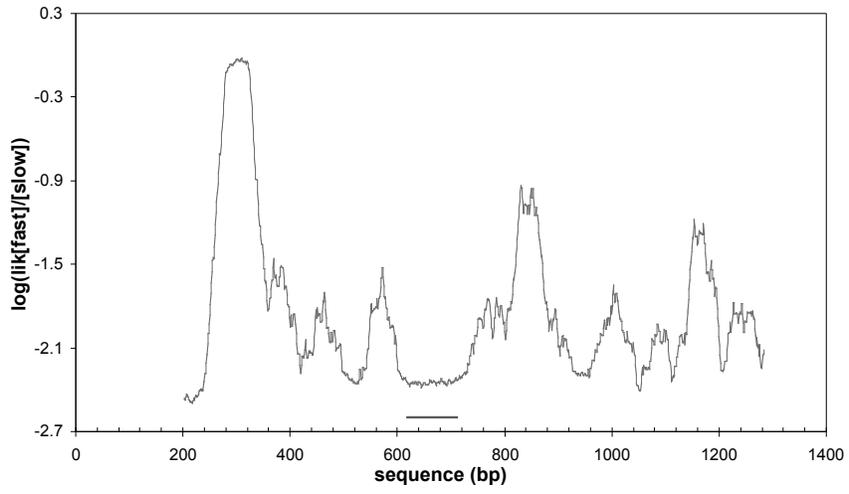
2. TEST-CASE:

**Identification of new regulatory
elements of a primate-specific gene**

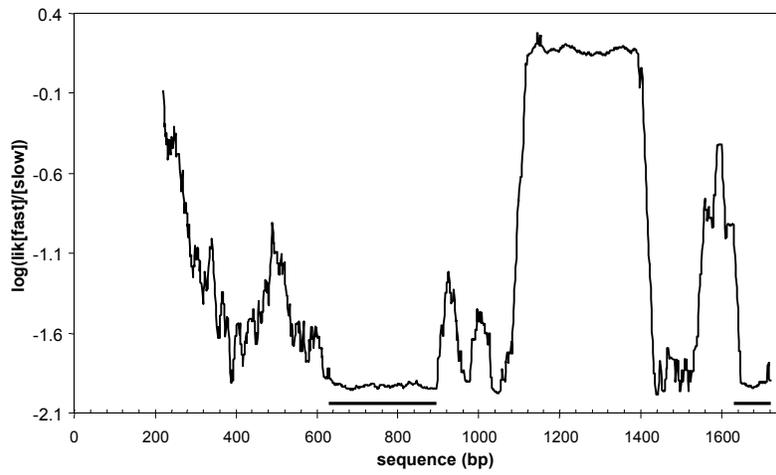
Apo B exon 19



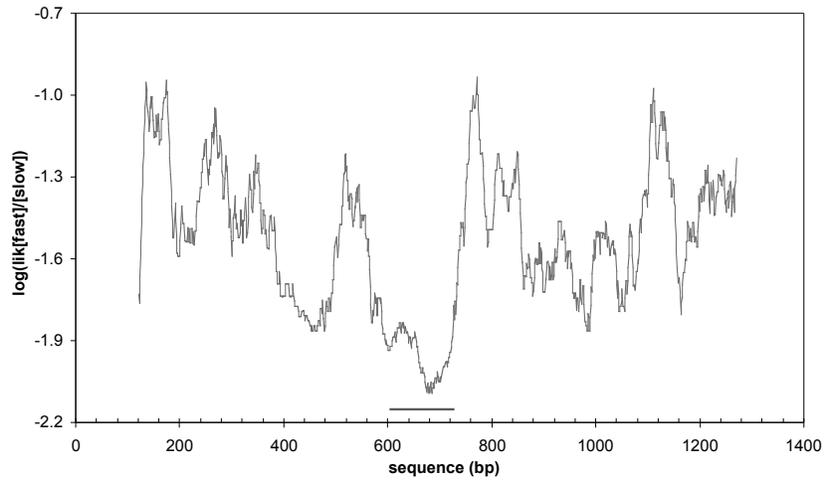
CETP exon 8



LXR-a exon 3

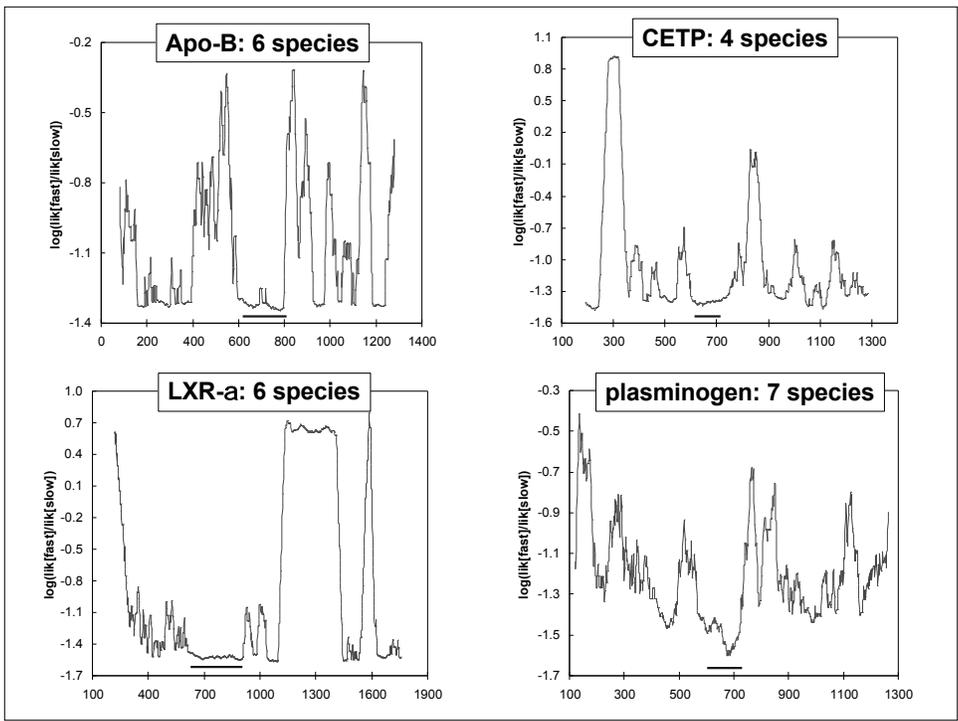
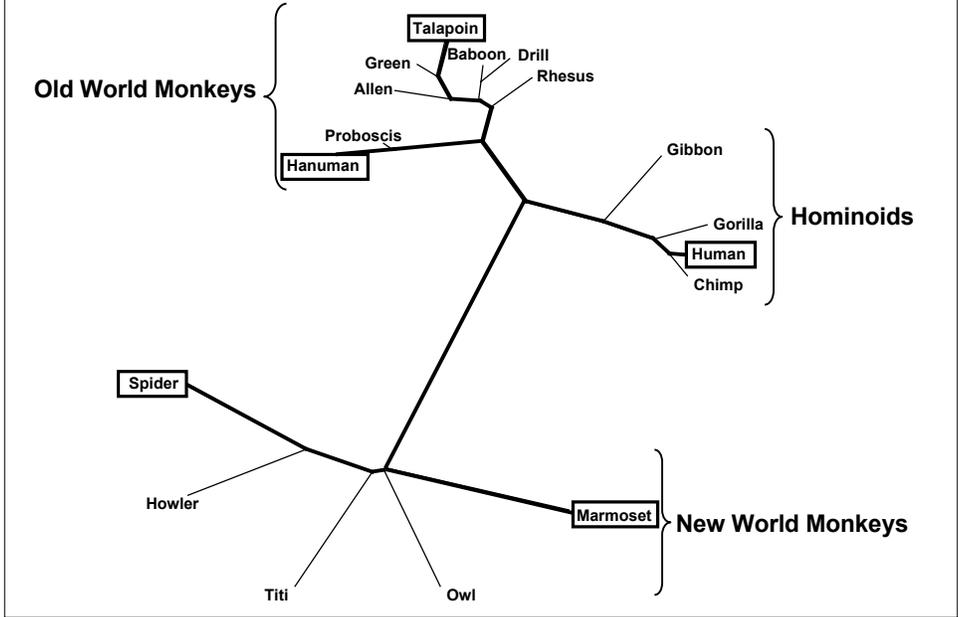


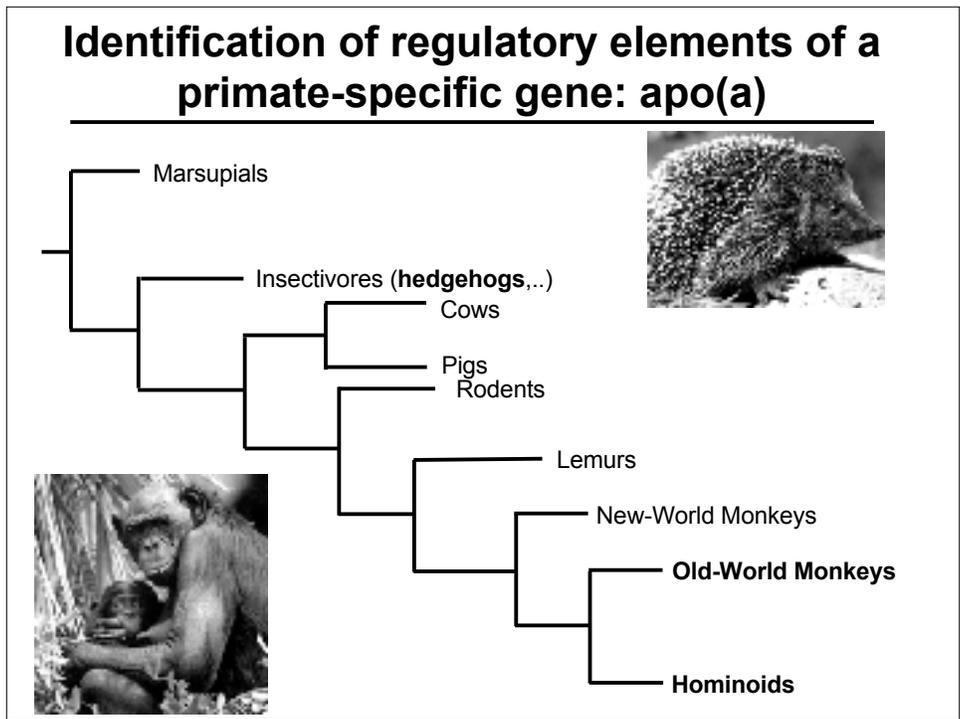
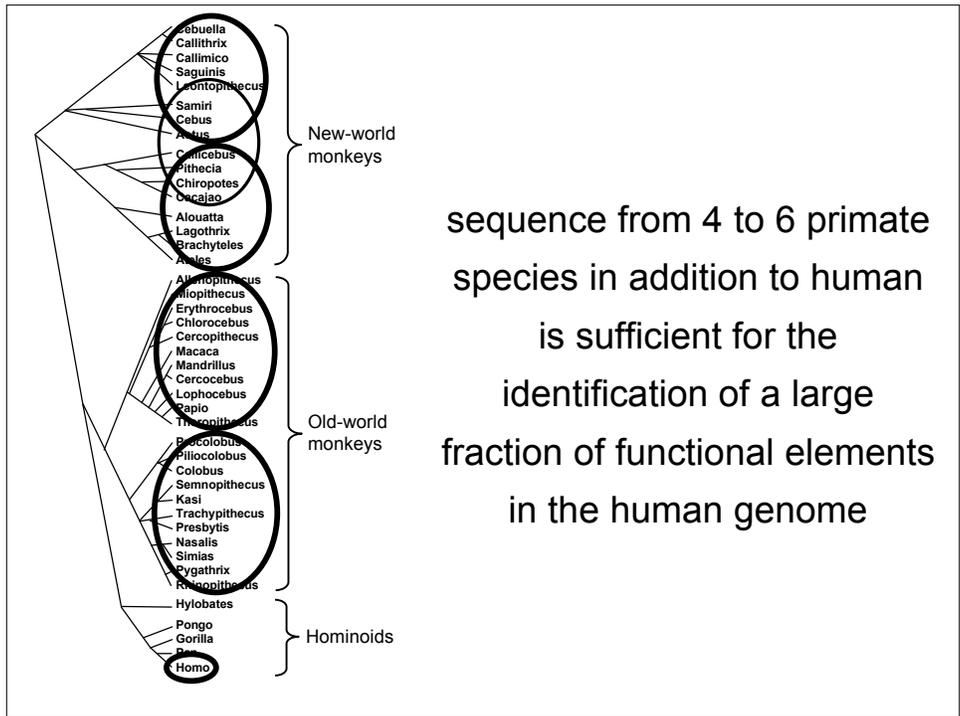
plasminogen exon 6



Amplification of evolutionary
distance through multiple species
comparison successfully identifies
known functional elements

Which primates should we sequence?

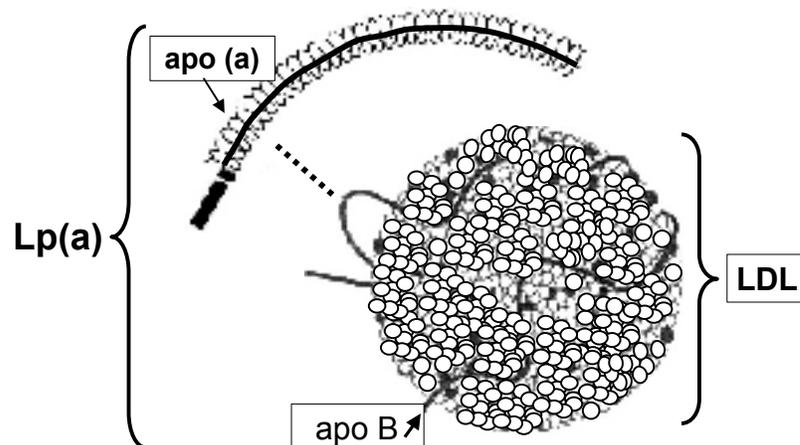




apolipoprotein(a)

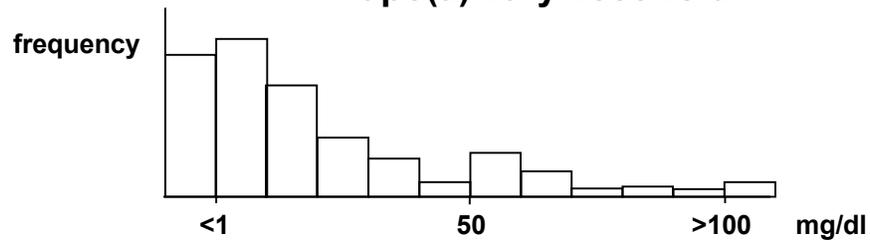
- Lacks a mouse ortholog
- People with high levels are at risk of developing atherosclerosis

apo (a) is a plasma protein



apo(a) plasma levels are highly variable

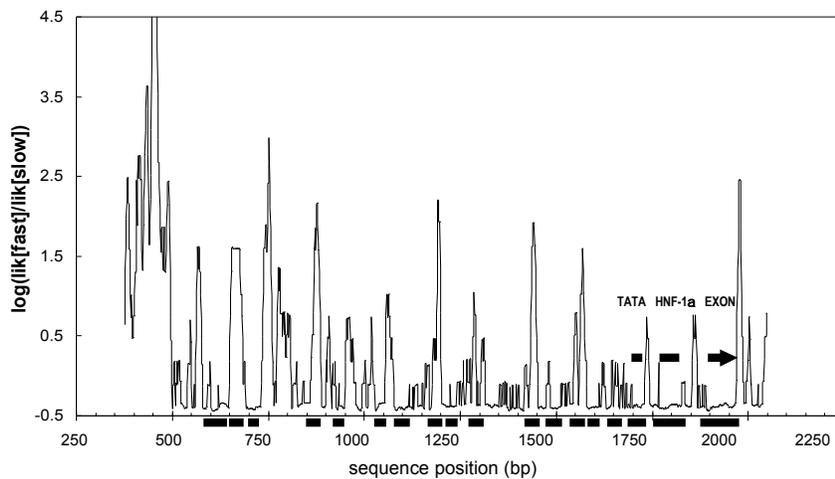
* apo(a) vary 1000-fold



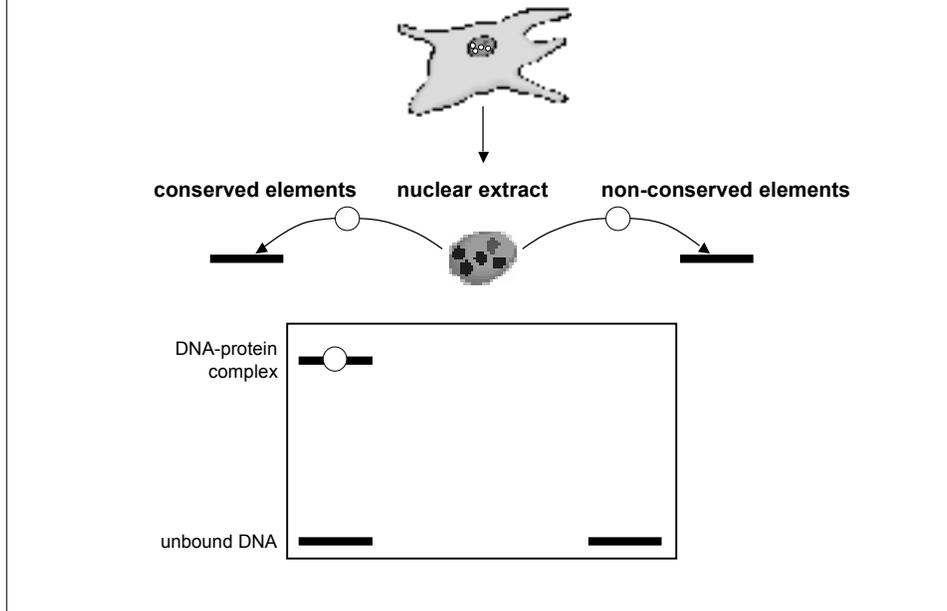
* standard protein ~10-fold

~ 50% of apo(a) variability in levels is attributable to the regulation of the transcription of the gene

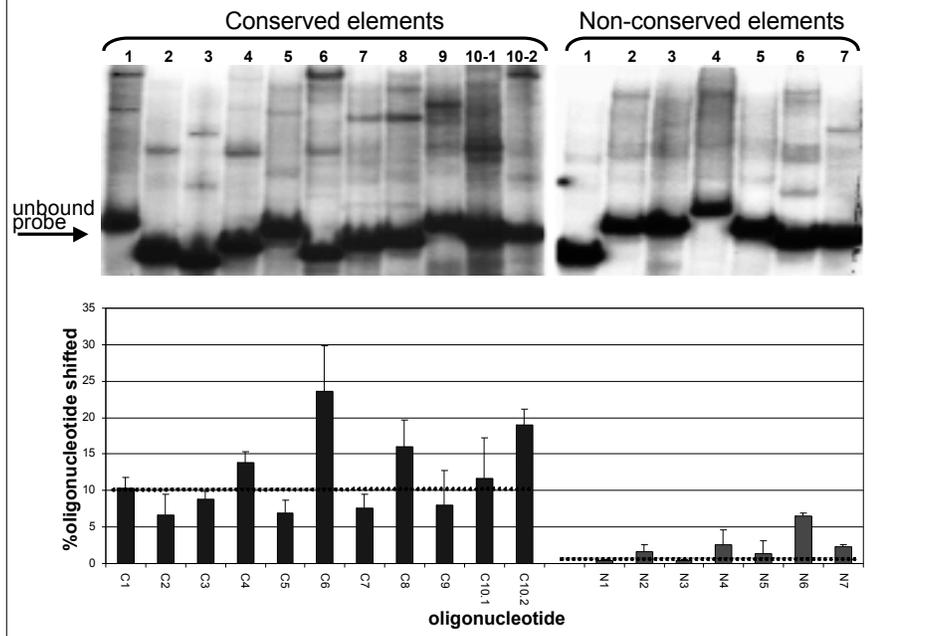
Phylogenetic shadowing of the apo(a) promoter



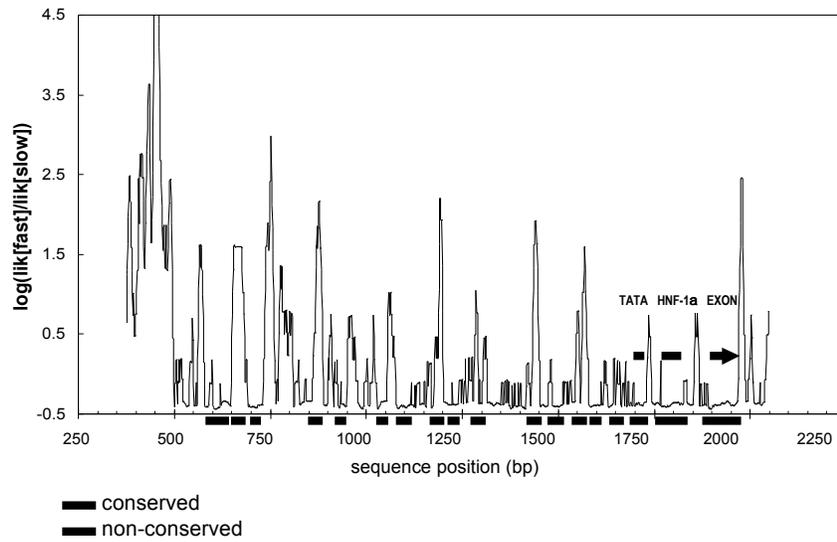
Gel-shift assay to assess DNA-protein interactions



Gel-shift analysis of conserved elements in the apo(a) promoter

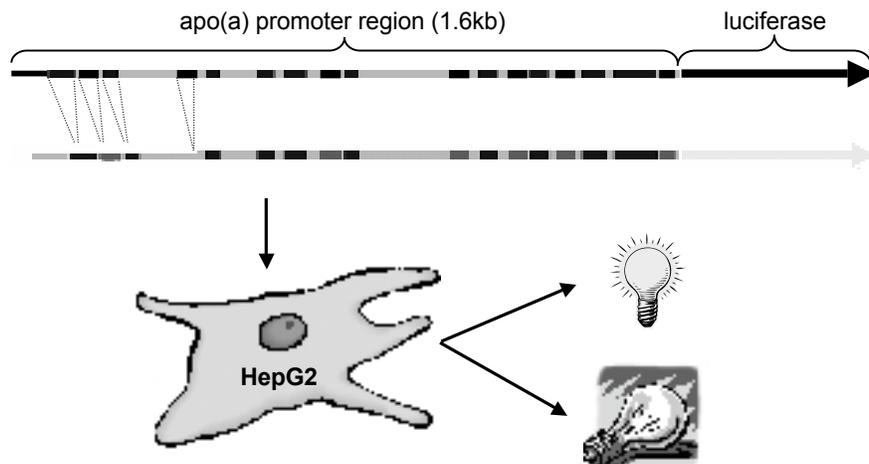


Phylogenetic shadowing of the apo(a) promoter

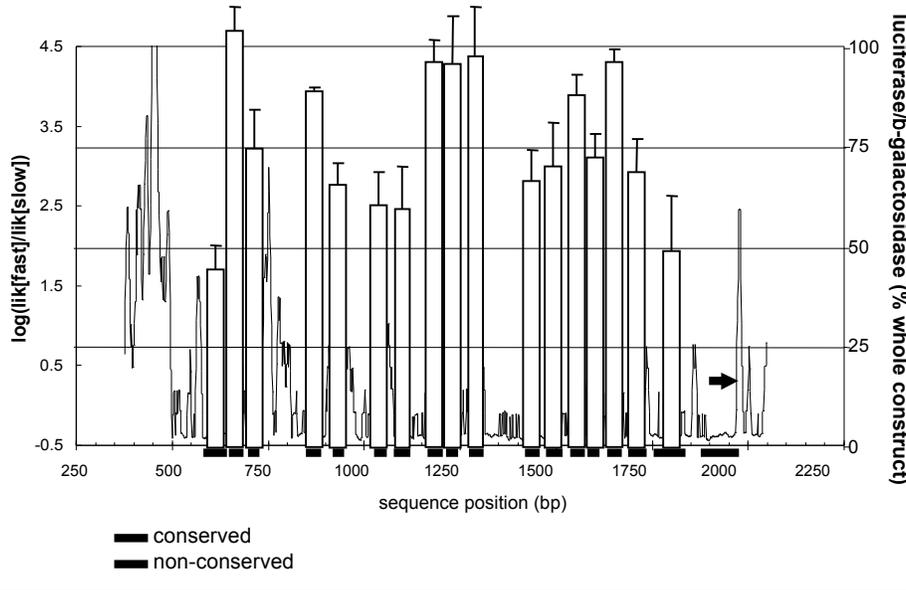


Testing the function of conserved elements

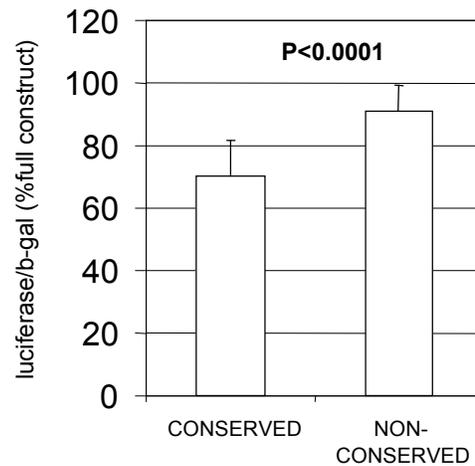
Cell culture transcription assay:



Phylogenetic shadowing of the apo(a) promoter

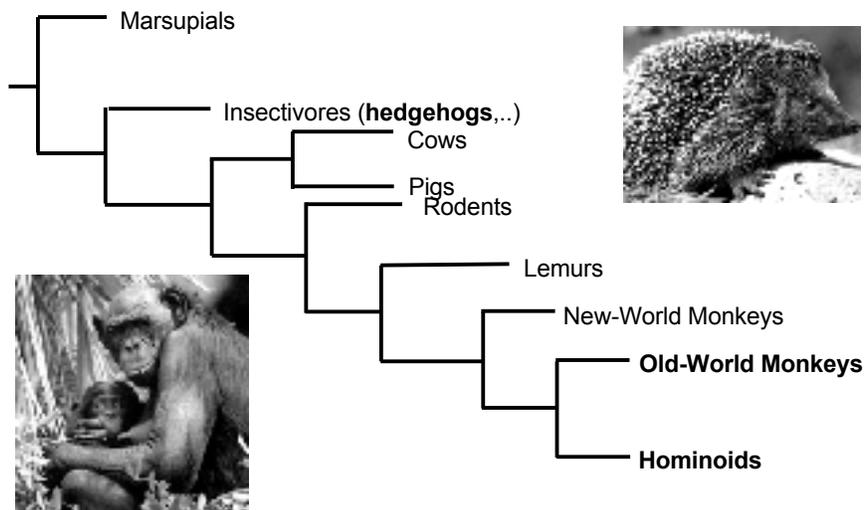


Average of expression data for conserved and non-conserved elements



genomic sequence from multiple primates can be used to identify functional sequence shared by this group of species

apo(a) is a recently evolved gene but...



gene loss **OR** convergent evolution?

Strategy to unravel the mechanism of evolution of the apo(a) gene

Sequence the genomic region containing apo(a) and its orthologous region in species without apo(a):

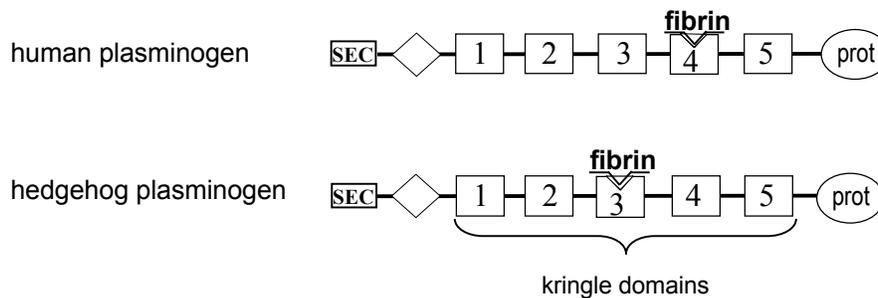
Species with apo(a):

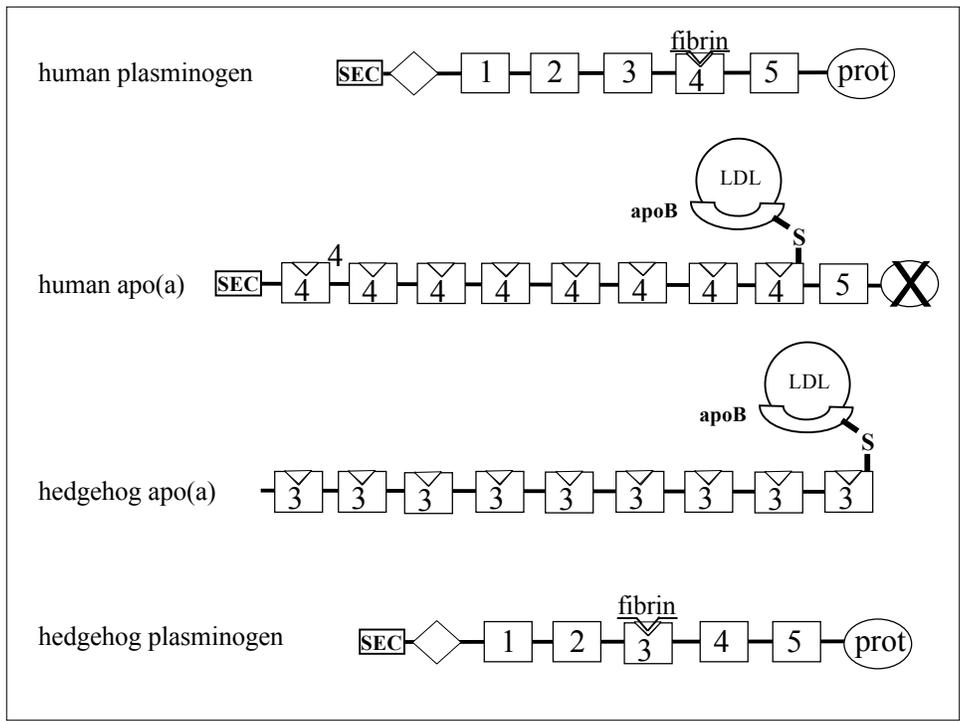
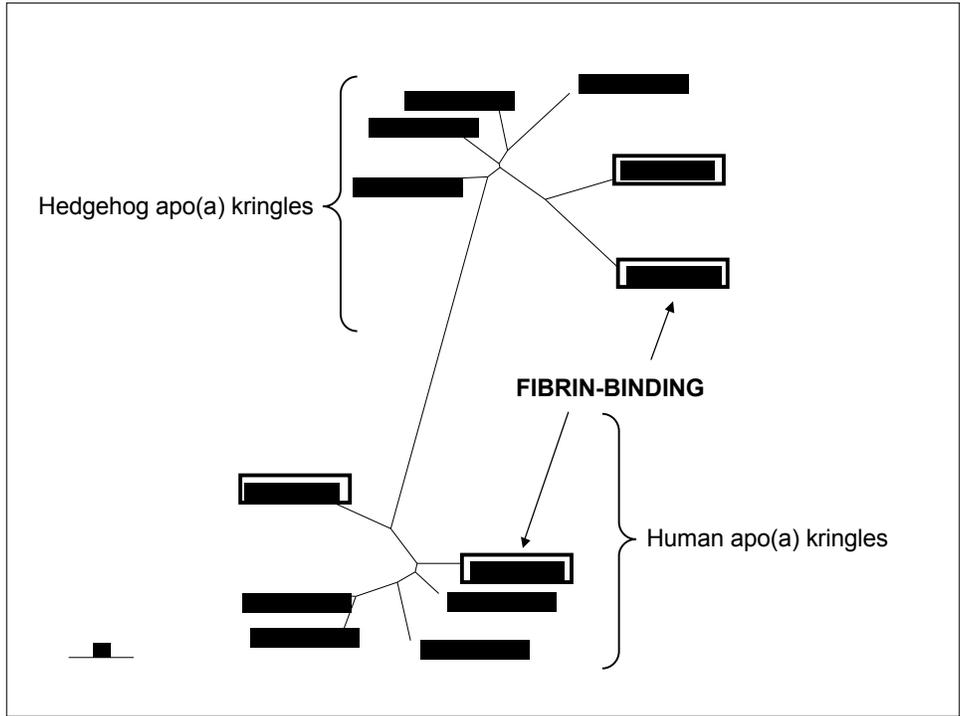
- human: publicly available
- baboon: sequence
- hedgehog: generate BAC library and sequence

Species without apo(a):

- mouse: sequence
- lemur: generate BAC library and sequence

human apo(a) is a duplication of plasminogen

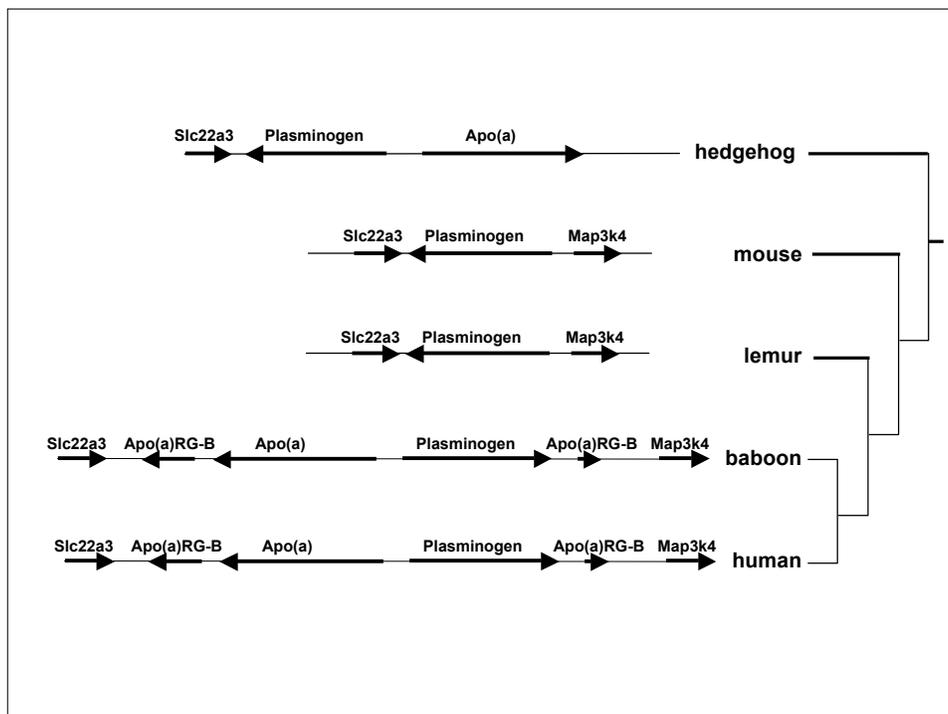




gene loss **OR** convergent evolution?

Independent remodeling of plasminogen:
CONVERGENT EVOLUTION

Can we detect apo(a) remnants in
species that do not express apo(a)?



apo(a) is a unique example of convergent evolution where the same gene arose twice through the independent duplication and remodeling of the same paralogous gene